

# Epidemioloogia IV. Epidemioloogiaajakirjade arvjoonised ja nende kvaliteet

Annika Voore<sup>1</sup>, Aleksei Baburin<sup>2</sup>, Allan Puur<sup>3</sup>, Mati Rahu<sup>2</sup>

Eesti Arst 2013;  
92(6):312–319

Saabunud toimetusse:  
09.09.2011  
Avaldamiseks vastu võetud:  
30.04.2013  
Avaldatud internetis:  
28.06.2013

<sup>1</sup> Crown CRO Oy,  
<sup>2</sup> Tervise Arengu Instituudi  
epidemioloogia ja  
biostatistika osakond,  
<sup>3</sup> Tallinna Ülikooli Eesti  
demograafia instituut

Kirjavahetajaautor:  
Mati Rahu  
[mati.rahu@tai.ee](mailto:mati.rahu@tai.ee)

Võtmesõnad:  
epidemioloogia, ajakirjad,  
arvjoonised, arvjooniseliigid,  
kvaliteet

**Eesmärk.** Teha kindlaks uurimistulemuste esitamisel kasutatavate arvjooniste ja arvjooniseliihide sagedus ning kvaliteet epidemioloogiaajakirjades.

**Meetodid.** Kolme epidemioloogiaajakirja (American Journal of Epidemiology, European Journal of Epidemiology, Paediatric and Perinatal Epidemiology) 2010. aasta numbritest valiti välja 187 algupäraste uuringute tulemusi kajastavat artiklit. Tehti kindlaks nendes esitatud arvjooniste ja arvjooniseliihide arv. Iga arvjoonise kvaliteeti – s.o visuaalset selgust, arvulist moonutatust, iseenesestmõistetavust, risustatust ja mõnda muud tunnust – hinnati R. J. Cooperi jt (2001, 2002, 2003) ning D. L. Schrigeri jt (2006) meetodika alusel.

**Tulemused.** Valitud artiklitest sisaldas arvjooniseid 50%, kokku oli avaldatud 184 arvjoonist. Sagedamini kasutati joondigrammi (38%), püst- ja rõhttulpdiagrammi (27%), punktidiagrammi (15%), elulemuskõverat ja hajuvusdiagrammi (kumbagi 5%). Üldse ei leidnud kasutamist karp- ega sektordigramm. Veatuid arvjooniseid oli 10%. Põhilised eksimused olid iseenesestmõistetavuse (56%) ja visuaalse selguse (42%) puudumine, arvuline moonutus (39%) ja liigsete andmete esitamine (27%).

**Järeldused.** Ehkki viimaste aastakümnete jooksul on tohutult paranenud andmetöötlusvõimalused ja uurimustes kasutatavate arvjooniste koostamise põhimõtteid käsitletud küllaldase selgusega, avaldatakse palju mannetuid jooniseid. Peame ülimalt soovitatavaks lisada arvjoonistele esitatavate üksikasjalike nõuete loetelu kõigisse juhendmaterjalidesse, mis puudutavad teadustöö tulemuste esitamist ajakirjades, raportites, magistri- ja doktoritöödes.

*Arvjoonis peab tööle, mille juurde ta kuulub, midagi lisama. Kui ta seda ei tee, pole teda ka vaja.*

Uno Mereste, Maimu Saarepera,  
1983 (1, lk 8)

Arvjoonis (*graph, chart, plot*) – joonis, millel arvandmeid kujutatakse geomeetriliste kujunditena (1) – on tavaline teaduskirjutise osa. Sarnaselt teiste teadusharudega kasutatakse arvjoonist tihti epidemioloogias. Nii sisaldavad umbes pooled ajakirja Epidemiology artiklid arvjooniseid (2). „Epidemioloogia entsüklopeedias“ (3) loetletakse arvjoonise üldiste oluliste omaduste seas „suure pildi“ näitamist (arvude sagedusjaotus, trend, võrdlused ja seosed), ekstreem- ja võõrväärtuste ning vigaste arvude hõlpsat avastamist ning mahuka teabe kajastamist väikesel pinnal.

Arvjooniseid on liigitatud mitmel alusel, näiteks joonise mahu (liht- ja liitdiagramm), geomeetrilise kujundi (joon-, tulp-, sektordigramm jt) ja selle mõõtme (ühe-, kahe- ja kolmemõõtmeline) ning kasutamiststarbe (struktuuri-, dünaamika-, võrdlusdiagramm jt) järgi (1). Seoses terviseandmete hulga suurenemise ja andmeanalüüsi meetodite arenemisega on arvjooniseliike aasta-aastalt mitmekesistatud ning täiustatud. Epidemioloogias on eraldi tähelepanu osutatud arvjoonistele, millel selgitatakse segava teguri (*confounding factor*) olemust (4–6), kõrvutatakse vähiregistrite andmeid riigiti (7), näidatakse haigestumuse ja suremuse ajatrende (8), käsitletakse põhjuslikke seoseid (9–11), teavitatakse haiguseriskidest (12) ning kujutatakse suhtelist riski ja šansisuhet (13).

Jooniseliigist olenemata peab teadustöös kasutatav teaduslik arvjoonis (*scientific*

graph) olema täpne, asjakohane ja üksnes vältimatult vajalike koostisosadega, vastama teatavatele esteetika- ja kompositsiooninõuetele (1, 14, 15). Paraku ei ole teaduskirjutiste autorid (ega lugejad) enamasti saanud vastavat koolitust, mistõttu piirduakse uurimustes mõne tuntuma arvjooniseliigi esitamisega ja nendegi koostamisel on tehtud mitmeid vigu (16).

Meie töö eesmärk oli kirjeldada uurimistulemuste esitamisel kasutatavate arvjooniste ja arvjooniseliikide sagedust ning kvaliteeti epidemioloogiaajakirjades. Artikkel moodustab osa Annika Voore 5. detsembril 2012 TÜ tervishoiu instituudis edukalt kaitstud magistritööst „Arvjooniste kasutamine epidemioloogias“.

## MEETODID

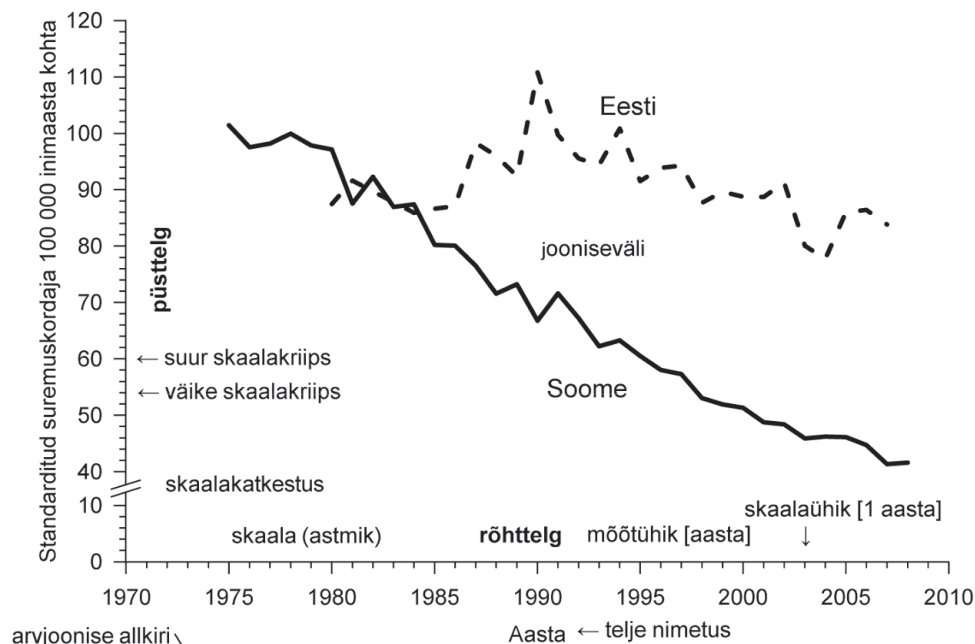
Ajakirja mõjuteguri (IF, *impact factor*) suuruse järgi (17) valisime välja kolm eri tasemega epidemioloogiaajakirja. Uuringu kavandamisel otsustasime varem üllitatud andmete alusel (18, 19), et tuleb läbi vaadata vähemalt 150 artiklit, s.o igast ajakirjast mitte alla 50, et saada analüüsimiseks kokku samapalju arvjooniseid. 150 arvjoonise korral kõigub mingi hinnatud tunnuse (nt arvjoonistel ilmnenu vea) suhtelise sageduse 95%

usaldusvahemik  $\pm 5\text{--}6\%$ , kui selle sageduse arvväärus on  $\leq 10\%$  või  $\geq 90\%$  (20).

Vajaliku arvu artiklite saamiseks valisime harvemini ilmuvatest või väiksema mahuga ajakirjadest kogu aastakäigu 2010: Paediatric and Perinatal Epidemiology (Paediatr Perinat Epidemiol; IF 2009 = 1,719) numbrid 1–6 ja European Journal of Epidemiology (Eur J Epidemiol; IF 2009 = 3,718) numbrid 1–12. Ajakirjast American Journal of Epidemiology (Am J Epidemiol; IF 2009 = 5,589) võtsime huupi (programm Research Randomizer (21)) 2010. aasta 171. köitest numbrid 2, 11 ja 12, ning 172. köitest 2, 8 ja 10.

Nimetatud ajakirjanumbritest trükkisime välja ainult algupäraste uuringute tulemusi kajastavad kirjutised (kokku 187), arvesse ei läinud juhtartiklid, ülevaateartiklid, metoodilised artiklid, juhised, kirjad ega kommentaarid. Seejärel analüüsisime arvjoonistega artikleid, mille hulka ei läinud need, mis sisaldasid ainult andmevoo skeeme (*data flow diagram*) või ainult muid skemaatilisi jooniseid. Tegime kindlaks artiklites sisalduvate arvjooniste arvu ja arvjooniseliigid.

Arvjoonise põhilised koostisosad on toodud näidisjoonisel (vt joonis 1). Samas ei ammenda viimane kõiki arvjoonistega



**Joonis.** Kopsuvähi standarditud suuremuskordajad (Euroopa standardrahvastik) Soome (1975–2008) ja Eesti (1980–2007) meesrahvastikus.

Arvandmed pärinevad Maailma Terviseorganisatsiooni andmebaasist (<http://data.euro.who.int/hfad>).

allmärkus

**Joonis 1.** Näidisjoonis joondigrammi põhiliste koostisosade kohta.

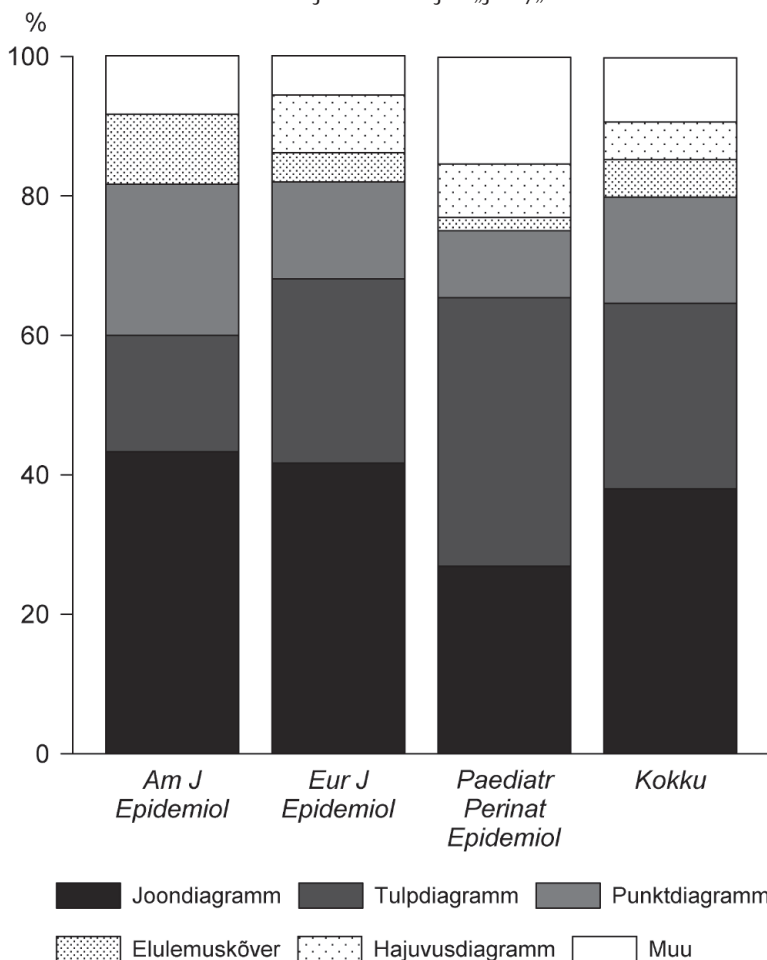
**Tabel 1.** Artiklite, arvjooniseid sisaldavate artiklite<sup>1</sup> ja arvjooniste arv uuringusse võetud 2010. aasta epidemioloogiaajakirjades

Ajakirja nimetus	Valitud artiklid		Arvjoonistega artiklid		Artiklite arv arvjooniste arvu järgi				Arvjooniste koguarv
	Arv	%	Arv	%	1	2	3	≥ 4	
Am J Epidemiol	58	100	30	51,7	14	8	5	3	60
Eur J Epidemiol	68	100	36	52,9	17	11	3	5	72
Paediatr Perinat Epidemiol	61	100	28	45,9	13	8	5	2	52
K o k k u	187	100	94	50,3	44	27	13	10	184

<sup>1</sup> Arvjoonistega artiklite hulka ei ole arvestatud neid, milles esinesid ainult vooludiagrammid või ainult muud skeemid. Selliseid artikleid oli igas ajakirjas kaheksa. Nimetatud jooniseliigid ei läinud arvjooniste arvestusse.

seotud oskussõnu, sest käsitleb üht jooniseliihi – joondidiagrammi.

Iga arvjoonise kvaliteeti hindasime Cooperi ja kaasautorite (18, 19, 22) ning Schrigeri ja kaasautorite (23) meetodika, ent käesoleva uuringu tarbeks mõnevõrra lihtsustatud kriteeriumide ja hindelehe järgi, milles jooniseosiste vastavust nõuetele väljendati kujul „jah“/„ei“.



**Joonis 2.** Arvjooniseliihikite sagedus uuringusse võetud 2010. aasta epidemioloogiaajakirjades. Kategooria „muu“ hõlmab histogrammi (5 joonist, 2,7%) ja ROC-kõverat (2 joonist, 1,1%) ning mõnda muud jooniseliihi. Jooniste hulgas ei olnud karpdiagrammi ega sektordiagrammi.

Arvjooniste arvustamisel läks puudusena kirja, kui

- sümbolid või lühendid (v.a üldkasutatavad) ei ole piisavalt defineeritud;
- joonis ei ole iseenesestmõistetav (joonise allkiri, legend jm ei sisalda kogu vajalikku teavet, et mõista kujutatut artiklit lugemata);
- jooniseosiste vahel esineb mingi vastuolu (nt legendi ja telje nimetuse mittevastavus);
- puudub visuaalne selgus (usaldusvahemikku tähistava joone otsakriipsude ebakorrapära; esitatakse liiga palju andmeid; kujundite kattumine; nimetused/selgitused on ebaselged);
- rakendatakse ebastandardset graafilist kujutusviisi (klassikalise arvjooniseliihi korral kasutatakse ebaharilikke osiseid, nt histogrammil esinevad tulpade vahel vahemikud; jooniseosised on jäänud selgitamata või on ebatüüpilised; usaldusvahemikku tähistavast joonest on toodud vaid pool);
- ilmneb arvuline moonutus (nt kujundite pikkus või pindala ei ole proportsionaalne andmetega; väär skaalakatkestus; skaalaühikud ei ole ühesugused; skaalakriipsud on ebaühtlased; asjatu 3D-joonis; jooned või tähipunktid on mõne rühma/nähtuse korral suuremad);
- joonise ja artikli (teksti, tabeli või teiste jooniste) andmed on vasturääkivad;
- oluline tunnus (segav tunnus, kihitamistunus) on jäänud joonisel kihitamata (nt joonisele ei ole kantud vajalikke andmeid eraldi meeste ja naiste, vanuserühmade, linna- ja maarahvastiku jms kohta);
- joonisel esitatakse liigseid andmeid (nt telje skaalal on vajalikud ühikud olemas, kuid need on toodud veel jooniseväljal; liigne teave legendis);
- joonis on risustatud (üleliigsed abijooned; mittevajalik lisatelg; kujundile lisatud

**Tabel 2.** Arvjooniste kvaliteedinõuete eiramise sagedus<sup>1</sup> uuringusse võetud 2010. aasta epidemioloogiaajakirjades

Kvaliteedinõude eiramine	Am J Epidemiol		Eur J Epidemiol		Paediatr Perinat Epidemiol		Kokku	
	Arv	%	Arv	%	Arv	%	Arv	%
<b>Arvjoonise osised</b>								
Arvjoonis ei ole iseenesestmõistetav	7	11,7	57	79,2	39	75,0	103	56,0
Visuaalse selguse puudumine	18	30,0	42	58,3	18	34,6	78	42,4
Arvuline moonutus	23	38,3	32	44,4	17	32,7	72	39,1
Ebastandardne graafiline kujutusviis	3	5,0	9	12,5	4	7,7	16	8,7
Vastuolud arvjoonise elementide vahel	6	10,0	7	9,7	1	1,9	14	7,6
Sümbolid või lühendid ei ole piisavalt defineeritud	0	0	9	12,5	4	7,7	13	7,1
<b>Arvjoonise seostatus artikli teksti/tabelitega</b>								
Arvjoonise ja teksti/tabeli / teiste arvjooniste andmete vasturääkivus	4	6,7	2	2,8	3	5,8	9	4,9
Oluline tunnus on jäänud arvjoonisel kihitamata	1	1,7	1	1,4	0	0	2	1,1
<b>Andmeesituse tõhusus</b>								
Arvjoonisel on esitatud liigseid andmeid	15	25,0	20	27,8	15	28,8	50	27,2
Arvjoonis on risustatud	0	0	21	29,2	6	11,5	27	14,7
Arvjoonis kajastab teavet, mis leidub tekstis või tabelis	3	5,0	1	1,4	0	0	4	2,2
Arvjooniste koguarv	60	100	72	100	52	100	184	100

<sup>1</sup> Eksimuste suhteline sagedus on arvutatud arvjooniste koguarvu suhtes. Et ühel ja samal joonisel esineb tihti mitu eksimust, siis ületab nende arv jooniste arvu.

vari; tausta asjatu toonitus; muaree-efekt);

k) joonisel kajastatakse arve/näitajaid, mis niigi esinevad tekstis või tabelites.

Kvaliteeti hindasid üheaegselt, kuid teineteisest sõltumatult, kaks isikut (AV, MR), kellest kumbki täitis hindelehe. Kui arvjoonise hindamise järel ilmnas, et hinnangud mingi tunnuse suhtes ei ühtinud, arutati kohe lahknemise põhjust ja jõuti edasise arutelu käigus üksmeelele.

Uuringu tulemused on väljendatud absoluutsete ja suhteliste sagedustena. Osa suhtarvude korral on esitatud veel 95% usaldusvahemik (uv), mis arvutati välja vabavaraprogrammiga (*Confidence interval for a proportion* (20)).

## TULEMUSED

Kolme epidemioloogiaajakirja 187 artiklist sisaldas arvjooniseid 50,3% (95% uv 42,9–57,6%) (vt tabel 1). Arvjooniste koguarv oli 184, seega tuli artikli kohta keskmiselt üks joonis. Arvjoonistega artiklitest oli ühe joonisega 46,8%, nelja või enama joonisega 10,6%.

Kõige enam kasutati joondigrammi (38,0%; 95% uv 31,0–45,5%), millele järgnesid püst- ja rõhttulpdiagramm (26,6%; 95% uv 20,4–33,6%) ning punktdiagramm (15,2%; 95% uv 10,4–21,2%). Nii elulemusköver kui ka hajuvusdiagramm moodustasid 5,5% joonistest. Ilmnas, et ajakirjas Paediatr Perinat Epidemiol eelistati tulpdiagrammi (38,5%), teistes ajakirjades joondigrammi (vt joonis 2). Üldse ei leidnud kasutamist karp- ega sektordiagramm.

Puudusteta arvjooniseid oli 10,3%, seejuures ajakirjas Am J Epidemiol 25,0%, Eur J Epidemiol 4,2% ja Paediatr Perinat Epidemiol 1,9%. Ühe-kahe veaga arvjooniseid esines 58,7% ja kolme või enama veaga 31,0%.

Tabelist 2 selgub, et väga sageli (56,0%; 95% uv 42,7–68,9%) avaldati arvjooniseid, mille igakülgne mõistmine osutub allkirja puudulikkuse tõttu võimatuks: lisaks tuleb lugeda artikli ülejäänud osi (teksti, tabeleid, teisi jooniseid). Kahes väiksema mõjuteguriga ajakirjas oli sedalaadi puudusega 77,4% (95% uv 62,6–89,9%) joonistest. Muude eksimuste seas tuli sagedamini ette

visuaalset ebaselgust, arvulist moonutust ja liigsete andmete paigutamist joonisele. Ajakirjade Eur J Epidemiol ja Paediatr Perinat Epidemiol arvjoonistele oli tihti kantud üleliigseid abijooni ning asjatult toonitatud tausta. Risustatus iseloomustas eriti Eur J Epidemiol jooniseid, seevastu ajakirja Am J Epidemiol joonised olid kujundatud asjatute abijooniteta ja muude mittevajalike osisteta.

## ARUTELU

Uuringust ilmnes, et epidemioloogiaartiklisse paigutati keskmiselt üks arvjoonis – valdavalt joon- või tulpdiaagramm, mis oli pahatihti vigane ega avanud ammendavalt käsitletava nähtuse sisu.

Kirjanduse andmetel ei erinenud teistel meditsiinierialadel arvjooniste keskmine arv artikli kohta kuigi palju meie leitud. See oli 0,9 ajakirjas Annals of Emergency Medicine (18) ja Journal of American Medicine (19).

Arvjooniseliikide kasutamise poolest sarnanesid epidemioloogiaajakirjad muude meditsiinierialade ajakirjadega: eelistati joon- ja tulpdiagramme. Need moodustasid kolmes anestesioloogiaajakirjas 85% (24), 11 mitme valdkonna ajakirjas 80% (16) ja ühes erakorralise meditsiini ajakirjas 51% (18) kõigist joonistest.

Kui asjakohane, oli joon- ja tulpdiaagrammil veel esitatud näitaja (valdavalt 95%) usaldusvahemik. Suhtelise riski (*relative risk*), šansisuhte (*odds ratio*) ja mõne muu ekspositsiooni efekti mõõtvana näitaja usaldusvahemiku eelistamine p-väärtusele on kujunenud epidemioloogias tavaliseks (25) ja juurdunud aegamööda arvjoonistel.

Teatavasti kipuvad algajad autorid liiga innustuma sektordiaagrammist, eriti selle värviküllasest 3D-teisendist. Kui keegi, olgu too juhendaja, retsensent või ajakirja toimetaja, kätt ette ei pane, raisatakse algeliste sektordiaagrammidega trükipinda ja jäetakse muud liiki arvjoonised teenimatult unarusse. Tõik, et 184 arvjoonise hulgas ei leidunud ühtki sektordiaagrammi, kinnitab soovitus – niisugused arvjoonised „kuuluvad kasutuimate jooniseliikide hulka ja neid tuleb teaduslikes käsikirjades vältida“ (26, lk 662) – omaksvõtmist. Samalaadset nõu on andnud veel Tufte (14) ja Freeman kaasautoritega (15).

Muidugi sõltub arvjooniste arv artiklis ja jooniseliigi eelistus oluliselt käsitletavast teemast või uuringukavandist. Viies tuntud üldmeditsiinilises teadusajakirjas (Annals of

Internal Medicine, British Medical Journal, Journal of American Medical Association, Lancet, New England Journal of Medicine) avaldati juhuslikustatud kliinilistele katsetele pühendatud kirjutistes keskmiselt 2,5 arvjoonist, kusjuures enim kasutati andmevooskeeme, Kaplani-Meieri elulemuskõvera (*Kaplan-Meier survival curve*) ja metaanalüüsis efekti tugevuse arvvärtuse ning selle usaldusvahemikuga diagramm (*forest plot, blobogram, Cochrane plot*) (27).

Kvaliteedi hindamisel kasutatud kriteeriumide järgi (vt eespool metoodika osas) osutus arvjoonistest veatuks kümnendik, madalaima mõjuteguriga ajakirjas vaid 2%. Ehkki hindamiskriteeriumide lahknemisest tingituna on eri autorite uuringutulemuste kõrvutamine raskendatud, toome siinkohal ikkagi ainsa kirjandusest leitud näitaja 11% (16).

Sagedaima veana (56%-l kõigist joonistest) **eirati iseenesestmõistetavuse nõuet**. Selle nõude kohaselt peab arvjoonisel kujutatatu olema mõistetav ilma artikli ülejäänud osi uurimata. Ehkki iseenesestmõistetavuse tagamine kuulub arvjooniste koostamise algtõdede hulka (28), ei ole hästi tuntud ajakirja joonistest ligi kolmandik (19) iseseisvalt kasutatav. Väga tihti sõnastatakse joonise allkiri ülilühidalt: unustatakse, et allkiri sisaldab teavet uuritava nähtuse, rahvastikurühma, piirkonna/riigi ja ajaperioodi kohta. Seega on allkirjad kujul „Uuritute vanusjaotus“, „Seos x-ekspositsiooni ja y-haiguse vahel“, „Z-haiguse levimus piirkonniti“, „Z1-haiguse suhteline risk vanuserühmiti“ jms teaduskirjutise jaoks vähehõlpsad ja kõlbmatud.

Teise sagedase eksimuse – **visuaalse selguse puudumise** – ilminguks on enamjaolt telgede nimetuse ärajätmine/ebatäielikkus/ohtrasõnalisus, samuti liiga paljude nähtuste üheaegne kajastamine jooniseväljal (joonis on muudetud ülikeeruliseks). Eriti riivab silma, kui skaala mõõtühikuks võetud protsendi korral ei asu sümbol „%“ mitte (ainult) skaalaotsa, vaid iga suure skaalakriipsu juures. Enamasti on tegu arvjooniste koostamise põhimõtete mittetundmise ja napi oskusega muuta joonist valmistava arvutiprogrammi vaikimisi seadistusi. Kui aga võrdluseks oleks mingi joonis, kus püsttelje mõõtühik ei ole mitte protsent, vaid „juhtu 100 000 inimaasta kohta“ – kas siis ikka sigineb tahtmine kanda see tekst iga (suurema) skaalakriipsu kõrval ...



Eksiti sellegagi, et viirutus- või toonitus-skaala ei olnud kooskõlas näitajate arvväär-tuste muutumisega ega taganud, et näiteks suuremale väärtusele vastab tumedam või tihedama viirutusega geomeetrilise kujundi pind.

**Arvulise moonutusega** oli 33–44% joonistest. Kirjanduse andmeil on selle vea sagedus 5–36% (19, 22). Analüüsimisel võetakse arvesse kõik eksimused, mis visua-liseerimisel ala- või ülehindavad kasutatud andmeid (18). Tihti jäetakse arvjoonise rõht- või püstskaala alguse väärtus märki-mata. Tüüpeksimuste hulka kuulub, et eri pikkusega skaalavahemikutele vastavad ühesu-guse pikkusega skaalavahemikud. Näiteks on kasutatud sama arvjoonise rõhtteljel võrdset skaalavahemikku nii poole kui ka ühe aasta pikkuse ajavahemiku (29) või teisel kord kolme, kord kuue kuu pikkuse vanusevahemiku tähistamiseks (30).

Siia eksimuste rühma kuuluvad 3D-joonised, mida taunitakse eeskätt põhjusel, et kolmanda mõõtme lisamine üksnes n-ö ilu pärast halvendab kujutatu mõistmist. Tulba, ringi või joone ruumili-seks muutmine ei aita üldjuhul arvjoonist paremini lugeda – pigem vastupidi – ega lisa uut teavet (15, 31).

Ka **koormatakse** arvjoonis üle **arvu-dega**, näiteks kirjutatakse tulpade või joone tähispunktide juurde arvväär-tused. Siinkohal loobume pikemast arutelust – Mereste ja Saarepera (1, lk 70) nendivad, et sedalaadi joonised näitavad “madalat arvjoonisekultuuri, oskamatus koostada joonist, mis oleks selge ja mõistetav ilma liigse arvukoormuseta”. Samade autorite arvates johtub joonisevälja üleküllastamine arvudega tabeli ja arvjoonise erinevuse mittemõistmisest.

Mõiste „**arvjoonise risustatus**“ (*chartjunk*) osutab asjaolule, et jooniseväljal paikneb liiga palju abijooni, esineb mittevajaalik lisatelg, taust on toonitatud või ilmneb muaree-efekt. Teinekord on risustatust määratletud laiemalt ja sellega hõlmatud veel kujundite kattumine, 3D-joonis või mingit üht nähtust/näitajat asjatult rõhutav värviskeem (22, 26).

Epidemioloogiaartiklites on suhtelise efekti mõõtmisel kinnistunud komme tähis-tada arvjoonisel abijoonega väärtus „1“, mis kajastab ekspositsiooni efekti puudumist. Veel ilmneb suunitlus mitte paigutada muid abijooni jooniseväljale: kui on olemas teljed,

nendel skaalad ja skaalakriipsud, näidatud mõõt- ja skaalavahemikud ning telgede nime-tused, on joonis loetav.

Omaette teema seondub muaree-efek-tiga – geomeetrilise kujundi mustri kirju-susest tekitatud optilise moonutusega, mis tugevasti häirib arvjoonise lugemist. USA statistik ja andmete visualiseerimise pioneere Edward Tufte on raamatus „The visual display of quantitative information“ (14) pühendanud peatüki joonise risusta-tusele, eriti aga opkunstist inspireeritud muustriliste arvjooniste mõttetusele. Paraku moodustas aastatel 1980–1982 muustriliste arvjooniste osakaal 11–21% tippajakirjades Nature, Lancet, Science ja New England Journal of Medicine. Mitmetes arvjooniseid käsitlevates õpikutes ja graafikapakettide kasutusjuhendites leidub näidisenäitena esitatud muustriga jooniseid enam kui muustrita – „praegu teeb arvuti paugupealt risustatud jooniseid“ (14, lk 111).

Stengel ja kaasautorid on osutanud, et tänapäeva nutikas tarkvara ahvatleb looma värvikirevaid ja seeläbi otsekui mõjuka-vamana näivaid arvjooniseid. Ent, nagu autorid on toonitanud, „kui joonis vajab tähelepanu köitmiseks [mitmeid] värve, on ta tõenäoselt kasutu. Teadlase kutsemeister-likkus avaldub mitte arvjooniste muutmises kunstiloominguks, vaid nende võimalikult lihtsas ja arusaadavas kujunduses.“ (26, lk 661). Loomulikult ei käi jutt vastuseisust mitmevärvitrükis informatiivsele arvjoonisele kui uurimisvahendile, vaid ikka valge-halli-musta skaala hülgamisest tava-pärase lihtsa joonise tegemisel. Teinekord võib koguni näha jooniseid, millele on nn graafilise disaini huvides jäetud joonestama mõni põhiosis, nt rõht- või püsttelg; seesugused skeemid on ristitud pseudoo-joonisteks (*pseudographs*) (22). (Graafiteoorias tähistab sama ingliskeelne termin teatud graafiliiiki – pseudograafi (32).)

Arvjooniste kvaliteeti on hinnatud suhte-liselt harva. Niisuguse hindamise üheks kitsaskohaks peetakse asjaolu (18), et vaat-luse ja seega võimaliku kriitika alt jäävad välja arvjoonisteta artiklid, mille analüüs võinuks osutada jooniste vajalikkusele. Teiseks, ehkki me tuginesime hindamisel varem kasutatud kriteeriumidele ja hinde-lehele, jääb eksimuste tuvastamine ikkagi mõnevõrra subjektiivseks. Joonise analüüsil me ei kasutanud pimemenetlust: kummagi hindaja käes oli artikli täielik koopia, autorid

<sup>1</sup> Crown CRO OY, Tartu, Estonia

<sup>2</sup> Department of Epidemiology and Biostatistics, National Institute for Health Development, Tallinn, Estonia

<sup>3</sup> Estonian Institute for Population Studies, Tallinn University, Tallinn, Estonia

Correspondence to:  
Mati Rahu  
mati.rahu@tai.ee

Key words:  
epidemiology, journals,  
graphs, graph types, quality

ja ajakiri olid teada. Hindajatena ei olnud me eelnevalt saanud formaalset arvjooniste eriõpet, hindamisoskused omandasime töö käigus ja vajaduse korral sirvisime ülivärtuslikku käsiraamatut (1).

Üldreeglina ei joonestata kirjastuses käsikirja jooniseid ümber. Teadusajakiri saab autorkonna graafilisi tõekspidamisi mõjutada autori- ja retsensendijuhendite kaudu. Kahjuks kasutavad ajakirjad seda hooba loiult ja juhendid sisaldavad arvjoonistele esitatavaid nõudeid napilt. 120 meditsiiniajakirjast 113 (94%) autorijuhendid ei anna nõu arvjooniste tegemiseks (33). Meie vaadeldud kolme epidemioloogiaajakirja autorijuhendites on toodud näpunäiteid jooniste konstrueerimise kohta, kõige põhjalikumalt on teinud seda Am J Epidemiol.

Ligi 30 aastat tagasi antud üldhinnangu järgi (1) iseloomustas Eesti trükistes esitatud arvjooniseid sageli küündimatu ülesehitus ja sisuline väheütlevus. Aga praegu? Arvjooniseid paigutatakse väga paljudesse uurimustesse, sh mitut laadi raportitesse, magistri- ja doktoritöösse. Arvjoonised moodustavad tihti lõviosa meie stendi- ja PowerPointi-ettekannetes. Mida me teame nende arvjooniste kvaliteedist?

Autoritena oleme veendunud, et käesolev artikkel annab õpialtile lugejale sõnumi – ärge korra arvjooniste koostamisel vigu, mida tehti vaatluse all olnud ajakirjades 90% jooniste korral. Heitke pilk artikli metoodikas loetletud 11 tüüpi vigadele, et haritud joonisetegijana neid vigu vältida.

## JÄRELDUSED

Ehkki viimaste aastakümnete jooksul on tohutult paranenud andmetöötlusvõimalused ja küllaldase selgusega käsitletud uurimustes kasutatavate arvjooniste koostamise põhimõtteid, avaldatakse palju mannetuid jooniseid. Peame ülimalt soovitatavaks lisada arvjoonistele esitatavate üksikasjalike nõuete loetelu kõigisse juhendmaterjalidesse, mis puudutavad teadustöö tulemuste esitamist ajakirjades, raportides, magistri- ja doktoritöös.

## TÄNUAVALDUS

Uuring on valminud Haridus- ja Teadusministeeriumi (SF0940026s07, SF0130018s11) ning Eesti Teadusagentuuri (IUT5-1) toel.

## HUVIKONFLIKT

Autorite kinnitusel puudub.

## SUMMARY

### Epidemiology IV. The frequency and quality of graphs in epidemiology journals

Annika Voore<sup>1</sup>, Aleksei Baburin<sup>2</sup>, Allan Puur<sup>3</sup>, Mati Rahu<sup>2</sup>

**OBJECTIVE.** To determine the frequency and quality of graphs used for presenting study findings in epidemiology journals.

**METHODS.** A total of 187 original research articles were selected from 2010 issues of *American Journal of Epidemiology*, *European Journal of Epidemiology*, and *Paediatric and Perinatal Epidemiology*. The articles were evaluated to determine the number of graphs and of graph types. The quality of each graph, i.e., visual clarity, numeric distortion, self-explanatoriness, chartjunk, and some other characteristics, was assessed using the methodical approach by R. J. Cooper et al (2001, 2002, 2003) and D. L. Schrager et al (2006).

**RESULTS.** Half of the articles contained graphs; the total number of graphs was 184. The five most common types of graph were the line graph (38%), the vertical and the horizontal bar graph (27%), the one-way graph (15%), the survival curve and the scatter graph (5% each). Neither the box-and-whisker graph nor the sector graph was used. Only 10% of the graphs did not include errors. The major errors were: the graph was not self-explanatory (56%), lack of visual clarity (42%), numeric distortion (39%) and over-abundance of data (27%).

**CONCLUSIONS.** Regardless of huge developments in data processing techniques in recent decades, and of established clear principles regarding the construction of graphs for research data, poorly drawn graphs are proliferating. We consider it to be highly desirable to include detailed requirements for graphs in all guidelines addressing the presentation of research results in journals and reports, as well as in master's and doctoral theses.

## KIRJANDUS/REFERENCES

- Mereste U, Saarepera M. Arvjoonised. Tallinn: Valgus; 1981.
- Wilcox AJ. Epidemiology and graphic design: new plans, and a new prize. *Epidemiology* 2001;12:477–8.
- Robbins NB. Graphical presentation of data. In: Boslaugh S, ed. *Encyclopedia of epidemiology*, 1. Los Angeles: Sage Publications; 2008; p. 443–6.
- Rothman KJ. A pictorial representation of confounding in epidemiologic studies. *Chronic Dis* 1975;28:101–8.
- Flanders WD, Johnson CY, Howards PP, Greenland S. Dependence of confounding on the target population: a modification of causal graphs to account for co-action. *Ann Epidemiol* 2011;21:698–705.
- Howards PP, Schisterman EF, Poole C, Kaufman JS, Weinberg CR. „Toward a clearer definition of confounding” revisited with directed acyclic graphs. *Am J Epidemiol* 2012;176:506–11.
- Rahu M. Graphical representation of cancer incidence data: Chernoff faces. *Int J Epidemiol* 1989;18:763–7.
- Devesa SS, Donaldson J, Fears T. Graphical presentation of trends in rates. *Am J Epidemiol* 1995;141:300–4.
- Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37–48.
- Hernán MA, Cole SR. Causal diagrams and measurement bias. *Am J Epidemiol* 2009;170:959–62.
- Joffe M, Gambhir M, Chadeau-Hyam M, Vineis P. Causal diagrams in systems epidemiology. *Emerg Themes Epidemiol* 2012;19:1.
- Ancker JS, Chan C, Kukafka R. Interactive graphics for expressing health risks: development and qualitative evaluation. *J Health Commun* 2009;14:461–75.
- Levine MA, El-Nahas AI, Asa B. Relative risk and odds ratio data are still portrayed with inappropriate scales in the medical literature. *Clin Epidemiol* 2010;63:1045–7.
- Tufte ER. *The visual display of quantitative information*. 2nd ed. Cheshire: Graphics Press; 2001.
- Freeman JV, Walters SJ, Campbell, MJ. *How to display data*. Oxford: Blackwell Publishing; 2008.
- Salmelin R. Graphical representation of statistical results in medical research. *Acta Universitatis Tampereensis* 551. Tampere: University of Tampere; 1997.
- Thomson Reuters. Web of Science. [http://thomsonreuters.com/products\\_services/science/science\\_products/a-z/web\\_of\\_science](http://thomsonreuters.com/products_services/science/science_products/a-z/web_of_science).
- Cooper RJ, Schriger DL, Tashman DA. An evaluation of the graphical literacy of *Annals of Emergency Medicine*. *Ann Emerg Med* 2001;37:13–9.
- Cooper RJ, Schriger DL, Close RJ. Graphical literacy: the quality of graphs in a large-circulation journal. *Ann Emerg Med* 2002;40:317–22.
- Confidence intervals for a proportion. Kevin's Web Page. Kevin Sullivan, Department of Epidemiology & Global Health, Emory University, Atlanta. <http://www.sph.emory.edu/~cdckms>.
- Urbanik GC, Plous S. Research Randomizer. Social Psychology Network, Wesleyan University, Middletown, CT, USA. <http://www.randomizer.org>.
- Cooper RJ, Schriger DL, Wallace RC, Mikulich VJ, Wilkes MS. The quantity and quality of scientific graphs in pharmaceutical advertisements. *J Gen Intern Med* 2003;18:294–7.
- Schriger DL, Sinha R, Schroter S, Liu PY, Altman DG. From submission to publication: a retrospective review of the tables and figures in a cohort of randomized controlled trials submitted to the *British Medical Journal*. *Ann Emerg Med* 2006;48:750–6.
- De Amici D, Klersy C, Tinelli C. Graphic data representation in anaesthesiological journals: a proposed methodology for assessment of appropriateness. *Anaesth Intensive Care* 1997;25:659–64.
- Rothman KJ, Greenland S, Lash TL. Precision and statistics in epidemiologic studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern epidemiology*. Philadelphia: Lippincott Williams & Wilkins, 2008:148–67.
- Stengel D, Calori GM, Giannoudis PV. Graphical data presentation. *Injury* 2008;39:659–65.
- Pocock SJ, Trivison TG, Wruck LM. How to interpret figures in reports of clinical trials. *BMJ* 2008;336:1166–9.
- Cox DR. Some remarks on the role in statistics of graphical methods. *Appl Stat* 1978;27:4–9.
- Ertel KA, Koenen KC, Rich-Edwards JW, Gillman MW. Antenatal and postpartum depressive symptoms are differentially associated with early childhood weight and adiposity. *Paediatr Perinat Epidemiol* 2010;24:179–89.
- Almqvist C, Garden F, Kemp AS. Effects of early cat or dog ownership on sensitisation and asthma in a high-risk cohort without disease-related modification of exposure. *Paediatr Perinat Epidemiol* 2010;24:171–8.
- Few S. *Show me the numbers: designing tables and graphs to enlighten*. Oakland: Analytics Press; 2004.
- Littover M, koost. Graafiteooria sõnastik – Graph theory glossary. Juhtimissüsteemide osakond, TTÜ Küberneetika Instituut: 2002. [http://www.cc.ioc.ee/jus/gtglossary/gtglossary\\_est.htm](http://www.cc.ioc.ee/jus/gtglossary/gtglossary_est.htm)
- Puhan MA, ter Riet G, Eichler K, Steurer J, Bachmann LM. More medical journals should inform their contributors about three key principles of graph construction. *J Clin Epidemiol* 2006;59:1017–22.

## TÜ Kliinikumi preemia parimale Eesti Arstis 2012. aastal ilmunud teadusartiklile

TÜ Kliinikum tunnustas juba kolmandat korda 1600eurose preemiaga möödunud kalendriaastal ajakirjas Eesti Arst ilmunud ja ajakirja toimetuskolleegiumi poolt parimaks peetud teadusartikli autoreid. Statuudi kohaselt esitati toimetuskolleegiumile preemia kandidaadid uuringute, ülevaadete ja haigusjuhu kirjelduste seast, mille hulgest tunnustati parimaks **Merike Jõgiste, Maire**

**Lubi, Aleksander Sipria ja Tarvo Rajasalu** artikkel „Mitmepalgeline feokromotsütoom: raske südamepuudulikkusega avaldunud haigusjuhu kirjeldus ja kirjanduse ülevaade“, mis ilmus Eesti Arsti 2012. aasta oktoobrinumbris, vt <http://www.eestiarst.ee/static/files/011/ea12111k556-564.pdf>. Preemia anti autoritele üle TÜ Kliinikumi kevadkonverentsil 24. mail 2013.